Implementation of stochastic multi attribute analysis (SMAA) in comparative environmental assessments

Prado, Valentina; Heijungs, Reinout

published in
Environmental Modelling and Software
2018

DOI (link to publisher)
10.1016/j.envsoft.2018.08.021

Environmental Modelling and Software 109 (2018) 223–231

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Implementation of stochastic multi attribute analysis (SMAA) in comparative environmental assessments

Valentina Prado[a,b,*], Reinout Heijungs[a,c]

[a] Institute of Environmental Sciences (CML), Department of Industrial Ecology, Leiden University, Einsteinweg 2, 2333 CC Leiden, The Netherlands
[b] EarthShift Global LLC, 37 Route 236, Suite 112, Kittery, ME 03904, United States
[c] Department of Econometrics and Operations Research, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081HV Amsterdam, The Netherlands

ARTICLE INFO

Keywords:
Comparative environmental analysis
Stochastic multi attribute analysis (SMAA)
Interpretation
Uncertainty

ABSTRACT

The selection of an alternative based on the results of a comparative environmental assessment such as life cycle assessment (LCA), environmental input-output analysis (EIOA) or integrated assessment modelling (IAM) is challenging because most of the times there is no single best option. Most comparative cases contain trade-offs between environmental criteria, uncertainty in the performances and multiple diverse values from decision makers. To circumvent these challenges, a method from decision analysis, namely stochastic multi attribute analysis (SMAA), has been proposed instead. SMAA performs aggregation that is partially compensatory (hence, closer to a strong sustainability perspective), incorporates performance uncertainty in the assessment, is free from external normalization references and allows for uncertainties in decision maker preferences. This paper presents a thorough introduction of SMAA for environmental decision-support, provides the mathematical fundamentals and offers an Excel platform for easy implementation and access.

1. Introduction

Quantitative environmental assessments such as life cycle assessment (LCA), environmental input-output analysis (EIOA) and integrated assessment modelling (IAM) can be used to inform environmentally motivated choices by government, industry and consumers (Arvesen et al., 2018; Castellani et al., 2016; Cellura et al., 2013; Gagnon et al., 2002; Groen et al., 2014; Hellweg and Mila i Canals, 2014; Hertwich et al., 2000; Pauliuk et al., 2017). These assessments are a particular type of environmental assessment, where the environmental impact is life cycle based, meaning it considers the systemic, cumulative impacts that a product, technology or service has in the environment. Results are based on a so-called functional unit which can vary in scale (from a consumer product, to a national wide policy implementation) and they quantify environmental impacts according to some already established set of indicators representing impact categories, such as global warming, ecotoxicity and acidification. However, while life cycle based assessments provide insight about broad environmental implications of a system, product or policy, results with several environmental indicators are difficult to interpret because there usually is no single best alternative (Finnveden et al., 2009; Laurin et al., 2016). Decision makers are then left to compare alternatives side by side across environmental indicators. Several environmental indicators alone can leave decision makers and practitioners alike subject to cognitive biases when determining the most preferable alternative (Hertwich and Hammitt, 2001). In the presence of trade-offs, studies often tend to narrow the analysis to a single environmental indicator (such as a carbon footprint), leaving decision makers to interpret several environmental indicators with no guidance, or generate a single score via a weighted sum with either ad-hoc or generic weight factors.

Calculation of an overall measure of environmental performance via a weighted sum W for each specific product alternative j, thus introducing W_j by Equation (1).

$$W_j = \sum_{i=1}^{m} w_i h_{ij} \quad (j = 1, \ldots, n) \tag{1}$$

where w_i is the weight of criterion (impact category) i and h_{ij} is product j's indicator value for that criterion. The approach has received fundamental criticisms with respect to the possibility of compensation (Norris, 2001; Pollesch and Dale, 2015, 2016; Rowley et al., 2012), and there are several practical limitations in compiling external references which are needed to ensure congruence and commensurability (Heijungs et al., 2007; Prado et al., 2017; White and Carty, 2010).

* Corresponding author. Institute of Environmental Sciences (CML), Department of Industrial Ecology, Leiden University, Einsteinweg 2, 2333 CC Leiden, The Netherlands.

E-mail address: v.prado@cml.leidenuniv.nl (V. Prado).

https://doi.org/10.1016/j.envsoft.2018.08.021
Received 1 August 2017; Received in revised form 10 August 2018; Accepted 22 August 2018
Available online 27 August 2018
1364-8152/

Software information

Name of software or data set Stochastic multi attribute analysis
Developer and contact address, telephone, fax and email numbers Same as authors
Year first available 2018
Hardware required, software required Computer and Microsoft office license
Availability and cost Free of charge given MS office
Program size 1.61MB

Besides concerns of oversimplification and subjectivity of a single score, practical challenges to the weighted sum are mostly due to data gaps and data constraints in the compilation of the external reference (Kim et al., 2013; Lautier et al., 2010). Any underestimation or overestimation of the external reference will distort the scaled result and ultimately the weighted results (Heijungs et al., 2007). In LCA for instance, some external references (e.g., for climate change) are easier to compile than others. Toxicity references for instance, have been a difficulty in LCA practice where the reference suffers from data gaps (Aboussouan et al., 2004; Pizzol et al., 2011) and as a result, scaled results consistently highlight the same aspects across diverse applications (Castellani et al., 2016; Prado et al., 2017). Some argue that these practical challenges can be solved with more complete external references (Kim et al., 2013), while others argue that the problem of the weighted sum goes beyond data repair efforts (Cucurachi et al., 2017; Prado et al., 2017). The weighted sum applies a linear aggregation approach which is fully compensatory. This means, that as long as an alternative continues to improve on a single issue, it will continue to compensate, and make up for poor performances in other criteria indefinitely, a property which has been linked to a weak sustainability perspective (Munda and Nardo, 2009; Pollesch and Dale, 2015; Rowley et al., 2012).

In addition to compensation, others call for the importance of taking into account mutual differences between alternatives, which is not taken into account in a weighted sum. This is an issue because in a comparison it is the distinct aspects that influence the decision (Prado-Lopez et al., 2015). In the field of decision analysis this has been referred to as the *range sensitivity principle* (Fischer, 1995), which shows how our preferences for a particular aspect (or criterion) change depending on how different alternatives perform in that aspect. For example, when evaluating the options of an important purchasing decision, price may be an important factor, but if options are very similar in price, price becomes less important for that particular decision – no point at bickering for the price, if all options cost practically the same. The aim is being able to incorporate such mechanisms of decision making in the decision support for environmental problems. For life cycle based assessments, there are multiple interpretation methods at the indicator level that evaluate mutual differences (Mendoza Beltran et al., 2018), but these do not necessarily lead to aggregation into a single score, and thus do not help resolve trade-offs between alternatives. In essence, current approaches of either a weighted sum, reducing the decision to a single indicator or applying analysis for each pair of alternatives at the indicator level, can unintentionally lead to burden shifting and/or suboptimal decisions.

Summing up, while there is valuable insight to be gained from quantifying a range of environmental impacts, the complexity of results hinders the capabilities to inform sensible decision making. To resolve these issues, some researchers have implemented methods from the multi criteria decision analysis (MCDA) field that can avoid full compensation and take into account mutual differences (see Brüggemann and Patil (2011), Janssen (1992) and Munda (2008) for a general overview of MCDA approaches and ranking problems). Among these, outranking algorithms stand out due to the fact that they are partially compensatory, scale alternatives according to mutual differences, rely on predefined value functions, can handle qualitative and quantitative criteria, and have been shown to incorporate uncertainty in the results (Benoit and Rousseaux, 2003; Cinelli et al., 2014; Matarazzo et al., 2013; Prado et al., 2012; Rogers and Seager, 2009; Rowley et al., 2012). Although some authors in the literature call for case specific value functions in environmental problems (Reichert et al., 2007), for the case of LCA type assessments this would be unsuitable given the varying scales of these assessments and the nature of environmental indicators used. For instance, there is no notion of how much should be the allowed cumulative ozone depletion impact of a load of laundry, a pair of jeans or a litre of biofuel. The information needed for creating value functions is not existent and attempting would make the assessment too laborious and out of reach for the majority of studies. Therefore, predefined value functions as in outranking are useful in the interpretation of comparative LCA type studies. Specifically, outranking based Stochastic Multi Attribute Analysis (SMAA) has been used in a number of comparative LCA studies as a way to aggregate results in a range of applications such as transportation fuels (Rogers and Seager, 2009), carbon nanotubes (Canis et al., 2010), detergents (Prado-Lopez et al., 2014), biofuel feedstocks (Rajagopalan et al., 2016), and photovoltaics (Ravikumar et al., 2018; Wender et al., 2014). Note that SMAA can also be used to refer to stochastic methods which may have different underlying aggregation algorithms such as SMAA-TOPSIS (Zhu et al., 2018), which is fully compensatory. Here, we focus the attention to a specific type of SMAA which combines a set of properties favourable for LCA type applications. SMAA has a potential to be applied in other types of environmental assessments and case studies, but the detailed working of SMAA for life cycle based assessments is not well described in literature, as most authors rely on implementations in specialised software.

This paper provides the mathematical foundations of how to implement SMAA in comparative multi impact quantitative environmental assessments with uncertain information, with an emphasis on life-cycle based analyses, such as LCA. For illustrative purposes, it contains a hypothetical example of three alternatives and three environmental indicators. We have added an Excel file as supplementary information in which all steps can be traced and modified when appropriate.

This paper is organized as follows. Section 2 presents the methodological approach starting with the hypothetical case study followed by a description of the calculation steps in SMAA. Section 3 shows the SMAA results, including intermediate results, as applied to the hypothetical case study. Section 4 discusses the meaning and implication of results and concludes the manuscript with closing remarks and areas for further research.

2. Methodology of SMAA

2.1. Outranking

Outranking refers to a family of methods to solve multi criteria decision analysis (MCDA) problems first developed by Bernard Roy (1985). Within outranking, there are different variations. ELECTRE (Elimination and Choice Expressing Reality) and PROMETHEE (Preference Ranking Organization METHod for Enrichment of Evaluations) are most widely used. PROMETHEE is considered to be easier to comprehend and perform (Behzadian et al., 2010; Benoit and Rousseaux, 2003; Figueira et al., 2016) and ELECTRE is often used for classification as opposed to ranking purposes (Domingues et al., 2015). Outranking methods can be described as decision aid methods rather than as decision analysis given their pragmatic approach to value functions (Hertwich and Hammitt, 2001; Tsoukiàs, 2008). Rather than constructing each value function on each criterion, demanding large cognitive efforts from stakeholders, outranking uses pair-wise comparisons and a predefined value function per criterion which facilitates its application in assessments over different scales (Benoit and Rousseaux, 2003; Rowley et al., 2012; Prado et al., 2012). This paper applies a stochastic version of PROMETHEE II that can generate an overall ranking of alternatives, taking into account uncertainty in the performances.

2.2. Hypothetical comparative LCA case study

To illustrate the idea of SMAA in environmental assessments, we apply a hypothetical comparative case study consisting of three alternatives (A, B, and C) and three environmental indicators (X, Y and Z). Distributions of environmental indicator results for each alternative for each indicator are defined as a probability density function (Fig. 1). The hypothetical example has been created in a way that each alternative is at least the best in one environmental aspect so to illustrate a difficult decision problem. From each distribution, a Monte Carlo sample is created with sample size $R = 1000$. Table 1 shows the first two Monte Carlo runs for the performances of alternatives A, B and C over environmental indicator X. Note the method is not limited to $R = 1000$ and the number of Monte Carlo runs can be increased as needed. Also note that for all criteria, given it is representative of environmental impact, a lower-is-better preference is applied.

2.3. Basic equations and example implementation

The following outranking steps are taken.

2.3.1. Defining a preference function

Outranking relies on pair-wise comparisons in each criterion (environmental indicator). First, d_{ijkr} corresponds to the difference between alternative pairs j and k on impact category i in Monte Carlo run r (Equation (2)). Table 2 shows sample calculation of for alternatives A, B and C for the first two Monte Carlo runs in environmental indicator X.

$$d_{ijkr} = h_{ijr} - h_{ikr} \quad (i = 1,\ldots,m;\ j,\ k = 1,\ldots,n;\ r = 1,\ldots,R) \tag{2}$$

The distinction between preference and indifference is made with a preference threshold (P_i) and an indifference threshold (Q_i) shown at the horizontal axis of Fig. 2. Preference thresholds can be derived from expert elicitation or from uncertainty in the performance. Following previous applications of SMAA in environmental problems, we derive preference thresholds from uncertainty in the performance as it is easier to apply in broader problems (Rogers and Bruen, 1998). Environmental assessments can pertain to scales at consumer product level where there is no notion of what could be considered an allowed difference in performance. In this respect, uncertainty based preference thresholds represent an advancement over previous applications (Tan, 2005), because it avoids subjective information at the scaling step and rather takes a more data exploratory approach (Brüggemann et al., 2008). Calculation of P_i and Q_i thresholds from the uncertainty aligns with previous efforts of distinguishability analysis for dealing with uncertainty in LCA (Basson and Petrie, 2007) (see Table 2).

Calculation of P_i and Q_i thresholds uses propagated uncertainty (Equations (3) and (4)), as in the estimated standard deviation of the data (s_{ij} as shown in Equations (5) and (6)). Results for the hypothetical case study are shown in Table 3.

$$P_i = \frac{-1}{n}\sum_{j=1}^{n} s_{ij} \quad (i = 1,\ldots,m) \tag{3}$$

$$Q_i = \frac{1}{2}P_i \quad (i = 1,\ldots,m) \tag{4}$$

$$s_{ij} = \sqrt{\frac{1}{R-1}\sum_{r=1}^{R}\left(h_{ijr} - \bar{h}_{ij}\right)^2} \quad (i = 1,\ldots,m;\ j = 1,\ldots,n) \tag{5}$$

$$\bar{h}_{ij} = \frac{1}{R}\sum_{r=1}^{R} h_{ijr} \quad (i = 1,\ldots,m;\ j = 1,\ldots,n) \tag{6}$$

The relative performance of alternatives is measured by the outranking score (vertical axis of Fig. 2), θ_{ijkr}, ranging from 0 to 1 and it is a function of the mutual difference of the pair of alternatives, d_{ijkr}, (Equation (7)).

$$\theta_{ijkr} = \begin{cases} 0 & \text{if } d_{ijkr} \geq Q_i \text{ (indifference)} \\ 1 & \text{if } d_{ijkr} \leq P_i \text{ (complete preference)} \\ \frac{Q_i - d_{ijkr}}{Q_i - P_i} & \text{if } P_i < d_{ijkr} < Q_i \text{ (partial preference)} \end{cases} \quad (i = 1,\ldots,m;\ j,\ k = 1,\ldots,n;\ r = 1,\ldots,R)) \tag{7}$$

The fact that the outranking score has a non-linear preference function with a limited range in outranking score implies that it avoids full compensation between criteria. This can help protect the outcome of results against extremes, where one exceptionally good performance

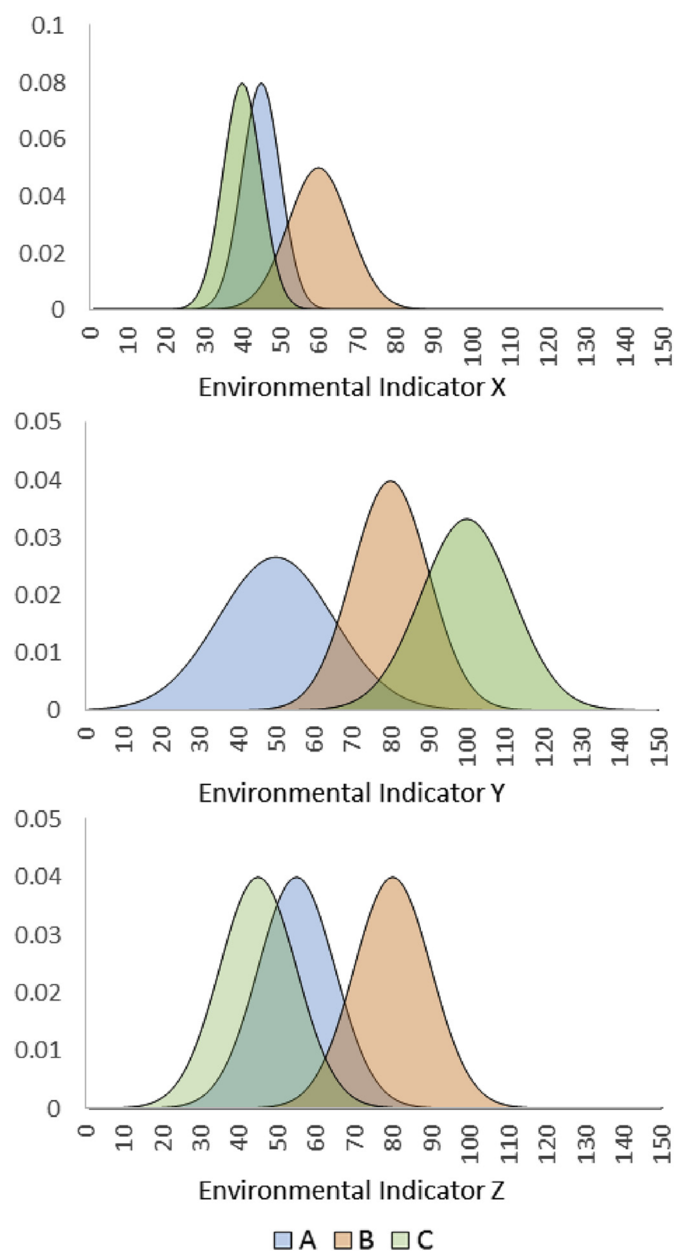

Fig. 1. Plots of probability density functions of results of alternatives A, B and C over environmental indicators X, Y and Z.

Table 1
Values of environmental indicator X for alternatives A, B and C in the first two Monte Carlo runs.

Alternative (j)	A	B	C
Run (r)	h_{XAr}	h_{XBr}	h_{XCr}
1	37.53	67.93	34.95
2	46.13	59.14	37.85
…	…	…	…

Table 2
Values of the pair-wise differences (both directions) of alternatives A, B and C in environmental indicator X in the first two Monte Carlo runs.

Alternative (j, k)	A-B, (B-A)	A-C, (C-A)	B-C, (C-B)
Run (r)	d_{XABr}, (d_{XBAr})	d_{XACr}, (d_{XCAr})	d_{XBCr}, (d_{XCBr})
1	− 30.40, (30.40)	2.58, (− 2.58)	32.98, (− 32.98)
2	− 13.02, (13.02)	8.28, (− 8.28)	21.30, (− 21.30)
…	…	…	…

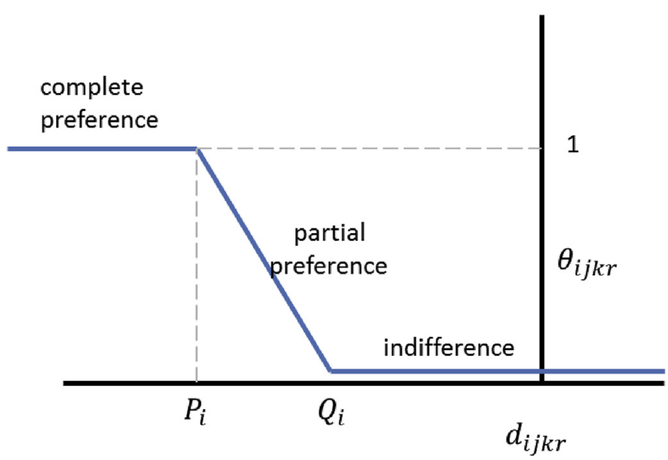

Fig. 2. Outranking preference function used in the case study where lower environmental impact is preferred.

Table 3
Preference and indifference threshold values for environmental indicators X, Y and Z.

Impact Category (i)	s_{iA}	s_{iB}	s_{iC}	P_i	Q_i
X	5.01	8.20	4.71	−5.97	−2.99
Y	15.54	9.68	12.06	− 12.43	− 6.21
Z	9.89	10.13	10.25	− 10.09	− 5.04

can compensate for poor performances in other areas and eventually dominate the final results (Stewart, 2008; Pollesch and Dale, 2015; Prado et al., 2017).

2.3.2. Calculating the outranking scores

The Monte Carlo ($r = 1, \ldots, R$) runs for all alternative ($j = 1, \ldots, n$) and all impact categories ($i = 1, \ldots, m$) yield values of the thresholds that define the preference functions in step a). A sample of results is shown in Table 4.

The relative performance of alternatives can be classified as:

- Complete preference: When there is enough evidence to determine that one alternative outperforms the other. The alternative obtains an outranking score of 1. As it is the case for alternative A as compared to B in the second MC run of environmental indicator X (Table 4, 2nd column, runs 1, 2 and 4).
- Partial preference: When there is a weaker preference between the pair of alternatives. The alternative will obtain an extrapolated outranking score between 0 and 1. For example alternative A as compared to B in run 5, achieves an outranking score of 0.86 because while it outranks B it does not reach full preference. When compared to alternative A, alternative B obtains a 0.
- Indifference: When the difference can be considered negligible, this is a “tie”, or the alternative is sufficiently outperformed by the other and the alternative obtains an outranking score of 0.For the first case, this can be observed in Table 4 in run 3 where alternatives all show a negligible differences with each other and they all obtain a 0 – a “triple tie”. Examples of an outranking score of 0 due to a “loss” (consequence of being compared to a superior alternative reaching *complete preference*) can be observed in run 4 where for each pair, there is a winning and losing alternative with outranking scores of 1 and 0 respectively.

In the MCDA literature the values θ_{ijkr} are known as positive and negative flows respectively (Behzadian et al., 2010). Positive when it indicates how much alternative j outranks alternative k and negative when it measures how much alternative j is outranked by alternative k. Note the term “flow”, refers to the outranking score of one alternative with respect to another per environmental indicator, not to material flows as it is commonly understood in the field of industrial ecology and related environmental analyses. (see Table 5).

2.3.3. Calculating the net flows

In every run, there will be an outranking score per pairwise comparison per environmental indicator. Over all runs, these result in the net flow, π_{ijr} (Equation (8)).

$$\pi_{ijr} = \sum_{\substack{k=1 \\ k\neq j}}^{n} (\theta_{ijkr} - \theta_{ikjr}) \quad (i = 1, \ldots, m;\ j = 1, \ldots, n;\ r = 1, \ldots, R) \tag{8}$$

2.3.4. Generating stochastic weights

In multi criteria decision methods, there are two types of subjective weights: trade-off weights and importance weights. Trade-off weights represent how much gains of a criterion (impact category) makes up for losses in another (Keeney, 2002). For instance, in a given problem, it could be that we are willing to increase impact by 5 kg of SO_2 eq, if that means a reduction of 2 kg of CO_2 eq - this would be the trade-off, the “even swap” (Hammond et al., 1999; Keeney, 2002). Importance weights on the other hand, reflect the relative importance of criteria according to the decision maker(s) values and do not depend on how alternatives perform. The application of either weight type depends on the preceding scaling method used. Outranking, in this case, the scaling method, applies importance weights (Riabacke et al., 2012).

Importance weights require preference information from either a panel or a decision maker. However, given that preference information is often unknown to stakeholders and analysts, we prefer in SMAA to use stochastic weights to reflect an inherent lack of knowledge about the weights. We define weight factors to have a range between 0 and

Table 4
Values of the pairwise outranking scores (both directions) of environmental indicator X in the first five Monte Carlo runs.

Alternative j, k, (k, j)	A vs. B, (B vs. A)	A vs. C, (C vs. A)	B vs. C, (C vs. B)
Run (r)	θ_{XABr}, (θ_{XBAr})	θ_{XACr}, (θ_{XCAr})	θ_{XBCr}, (θ_{XCBr})
1	1, (0)	0, (0)	0, (1)
2	1, (0)	0, (1)	0, (1)
3	0, (0)	0, (0)	0, (0)
4	1, (0)	0, (1)	0, (1)
5	0.86, (0)	0, (1)	0, (1)
…	…	…	…

Table 5
Values of the net flows for alternatives A, B and C over environmental indicator X in the first two Monte Carlo runs. Alternative C, being the one with the lowest impact in environmental indicator X, tends to have the lowest net score.

Alternative (j)	A	B	C
Run (r)	π_{XAr}	π_{XBr}	π_{XCr}
1	1	− 2	1
2	0	− 2	2
…	…	…	…

100, where the sum of weights equals 100. This information is defines the possible weight space for each criterion (Tervonen and Lahdelma, 2007). This approach has been proposed as a more inclusive alternative to the “equal weights” approach that limits evaluation of unknown weights to a single value (Rogers and Seager, 2009; Tylock et al., 2012). While single value weights may be defendable in those cases with knowledge of decision makers' preferences, it is not a robust approach in the absence of information regarding preferences. Stochastic weights, as applied in SMAA, assign a distribution to weight factor values. With stochastic weight factor values, the analyst obtains an aggregated overall score that takes into account all possible value systems, within reasonable bounds. Stochastic weights can also be used to reflect distinct preference information. Tylock et al. (2012) show how to modify weight factors according to different levels of importance in criteria (such as “below average”, “average” and “above average”). In this study, we limit the illustration to a situation where weight values are unknown and thus sampled equally for all impact categories. Previous research shows that the use of beta distributions, generates equally distributed weight values among criteria (Rogers, 2008). We apply the unconstrained weights calculation according to the pseudo Markov Chain by Tylock et al. (2012). Here, the first weight distribution is sampled from W_i (Equation (9)). Table 6 shows weight factor results over the first two Monte Carlo runs for the hypothetical case study.

$$W_i = \begin{cases} (100 - \sum_{t=1}^{i-1} W_t) \times beta\ (\alpha = 1,\ \beta = m - 1) & (i = 1, \dots, m - 1) \\ 100 - \sum_{t=1}^{i-1} W_t & (i = m) \end{cases} \tag{9}$$

The resulting weight distribution from the fixed value also resembles a beta distribution shape (Fig. 5).

2.3.5. Calculating overall scores

Aggregation of performances of alternatives across environmental indicators is done by a weighted sum of the net flows with the weight factors. Note that while the aggregated result can be calculated via a weighted sum, the scaling step in outranking in non-linear. The overall score, Z_{jr}, is obtained per run (Equation (10)). The overall score could be negative, but this is an indication of relative performance to other alternatives in the comparison (not of an absolute measure of impact or benefit to the environment) and it is used to rank alternatives where a higher overall score corresponds to a preferred performance. Table 7 shows the result of the first two Monte Carlo runs for the overall score of each alternative in the hypothetical case study.

$$z_{jr} = \sum_{i=1}^{m} w_{ir} \times \pi_{ijr} \quad (j = 1, \dots, n;\ r = 1, \dots, R) \tag{10}$$

2.3.6. Rank acceptability index

One way of communicating overall results, consolidated across all runs, is via the rank acceptability index which does a counting per run of the rank of each alternative. First, we rank the alternatives per run (Equation (11), Table 8).

$$\zeta_{jr} = \mathrm{Rank}(z_{1r}, \dots, z_{nr}, j) \quad (j = 1, \dots, n;\ r = 1, \dots, R) \tag{11}$$

Here, the symbol $\mathrm{Rank}(z_{1r}, \dots, z_{nr}, j)$ denotes the rank to be assigned to observation j in a series of observations $z_{1r}, \dots, z_{nr}$ (Lahdelma and Salminen, 2001).

Finally, a rank acceptability index assigns a probability value per alternative per rank (Lahdelma and Salminen, 2001). See Equation (12) for more details and Table 9 for an illustration.

$$R_{qj} = \frac{1}{R} \sum_{r=1}^{R} \delta(\zeta_{jr}, q) = \quad q = 1, \dots, n; j = 1, \dots, n \tag{12}$$

where $\delta(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & if\ x \neq y \end{cases}$ is the Kronecker delta.

If desired, this might be extended by a statistical test of equality of rank sums through the Kruskal-Wallis test (Cucurachi and Suh, 2015). However, in a context of Monte Carlo samples, null hypothesis tests should be treated with care (Heijungs et al., 2016). Therefore, we have decided not to include it in this exposition.

3. Results

3.1. Stochastic weights

Sampling $R = 1000$ values according to Equation (8) for $n = 3$ yields a weight factor distributions (Fig. 3). From this figure it can be observed that all weight factors value have approximately the same probability of occurrence in all three environmental indicators X, Y and Z. So indeed, this represents a default situation of equal weighting in the absence of real weights, but with a stochastic mindset.

3.2. Overall scores

After application of weight factors to the net flows we generate the overall scores, z_{jr}. Fig. 4 shows the distribution of z_{jr} for alternatives A, B and C. From this it can be seen that alternative A and C generate very similar scores and that alternative B has a relatively lower score. However, they appear to be competitive alternatives. These results are further analyzed via the rank acceptability analysis.

3.3. Rank acceptability index

An evaluation of the ranks in each run with the rank acceptability index (Eq. (12)) generates a probabilistic rank (Fig. 5). Similar to the spread of overall results, we see that the alternatives are quite competitive with each other. Alternative A and alternative C produce very similar results with a likelihood to rank first of 46% and 42% respectively. They also share similar rank acceptability indices for the second place (32% and 31%) and for the third place (23% and 27%). Alternative B shows a different profile and it is most likely a third best alternative (51%). Performance at the environmental indicator level (Fig. 1) shows that Alternative A has a relatively greater advantage in environmental indicator Y, while Alternative C holds a smaller advantage over environmental indicator X and Z. The eventual decision between A and C therefore comes down to the relative importance of indicators. Currently, all environmental indicators share the same priority. If a higher weight is placed in X and/or Z it would benefit alternative C, while if higher weight is placed in Y, then it would benefit

Table 6
Values of the stochastic weights for each impact category and the first two Monte Carlo runs.

Environmental Indicator (i)	X	Y	Z
Run (r)	w_{Xr}	w_{Yr}	w_{Zr}
1	1.94	51.11	46.95
2	16.40	8.32	75.28
…	…	…	…

Table 7
Weighted scores of the first two Monte Carlo runs.

Alternative (j)	A	B	C
Run (r)	z_{Ar}	z_{Br}	z_{Cr}
1	57.21	−50.83	−6.38
2	−58.64	33.31	25.33
...	...	...	...

Table 8
Ranked weighted scores of the first two Monte Carlo runs.

Alternative (j)	A	B	C
Run (r)	ζ_{Ar}	ζ_{Br}	ζ_{Cr}
1	1	3	2
2	3	1	2
...	...	...	...

Table 9
Ranked weighted scores for all three alternatives A, B, C.

Rank (q)	A	B	C
1	$\frac{456}{1000}$	$\frac{126}{1000}$	$\frac{418}{1000}$
2	$\frac{318}{1000}$	$\frac{367}{1000}$	$\frac{315}{1000}$
3	$\frac{226}{1000}$	$\frac{507}{1000}$	$\frac{267}{1000}$

alternative A.

4. Discussion and conclusion

This paper uses an example of a problem involving 3 alternatives and 3 environmental indicators to illustrate SMAA. Performances show a difficult selection as there is no single best alternative (Fig. 1). By applying SMAA and evaluating the mutual differences, it was possible to rank alternatives and while there is not a dominant single alternative, it was possible to identify a least preferable alternative (alternative B – see Fig. 5). To break the tie between alternatives A and C, it would be necessary to introduce weights that reflect a preference of an indicator over another, and/or refine performance data (Fig. 1). Therefore, given the possibility that this is an iterative approach, it could be that when faced with these results (a tie for the best alternative), the analyst invests research efforts into gathering specific weights or refines data concerning the underlying parameters affecting A's and C's performance in environmental indicators X and Z where they currently have very similar performances (Fig. 1). SMAA can be applied to other types of environmental assessment with quantitative uncertainty information and similar properties to life cycle based studies by means of the Excel tool provided. Extensions of the tool to accommodate more alternatives, indicators and/or Monte Carlo runs is straightforward by following the algorithms presented here. We recommend SMAA as a way to provide a sense of the overall relative ranking of alternatives as an additional way of evaluating results. This can help identify which alternatives are at the top or bottom based on overall performance in a way that it fulfills key conditions in environmental management (for an extensive review of the applicability of other MCDA approaches in sustainability refer to Cinelli et al., 2014).

The use of outranking in SMAA holds a strong sustainability perspective by limiting compensation with a nonlinear aggregation function. This will generate a more balanced outcome where a single good performance does not drive results (Pollesch and Dale, 2015; Prado et al., 2017). Moreover, value functions of outranking are applicable to any problem without the need of elicitation - a process known to be resource-intensive and ineffective depending on the scale of assessment (Kiker et al., 2005; Polatidis et al., 2006). For instance, what could be the allowed acidification potential of a kg of produce or a load of laundry? This process takes large cognitive efforts from stakeholders whom are not always available in environmental assessments.

Another key feature of SMAA as applied here is setting the preference thresholds in outranking as a function of the spread of the data to take into account mutual differences with respect to uncertainty. This means that the ultimate rank is sensitive to changes in uncertainty and hence responsive to changes in the quality of information. When the level of uncertainty changes, this will affect the calculation of the P and Q thresholds and the corresponding outranking scores, θ_{ijkr} (Eq. (7)). For example, when the standard deviation of alternatives A, B and C is twice as high (from 5 to 10, 8 to 16 and 5 to 10 respectively) in environmental indicator X, the rank acceptability indices respond so that alternative A has the highest likelihood of ranking first (49%) – Fig. 6. This illustrates the possible changes in the model in the event that new information comes along that increase the uncertainty propagated related to a particular environmental indicator. For the hypothetical case study, we see that when the difference between alternatives in environmental indicator X becomes less significant (same mean value, but larger dispersion), so does the preference for alternative C over alternative A. As a result, the likelihood of alternative A to rank first, increases.

Finally, SMAA is applicable under (high) uncertainty of performances and weights. With regards to weights, in the absence of real knowledge regarding preferences, we see sampling of the entire weight space as an improvement over discrete equal weights. When preferences are known, then we suggest applying the distinct preferences in a way that still accounts for those uncertainties (See Tylock et al., 2012 for an example of modified stochastic weights).

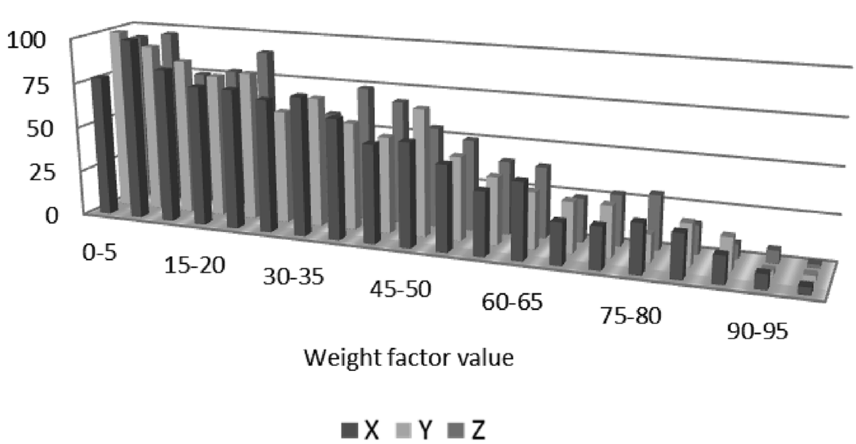

Fig. 3. Distribution of weight factors for impact categories X, Y and Z, based on 1000 Monte Carlo runs.

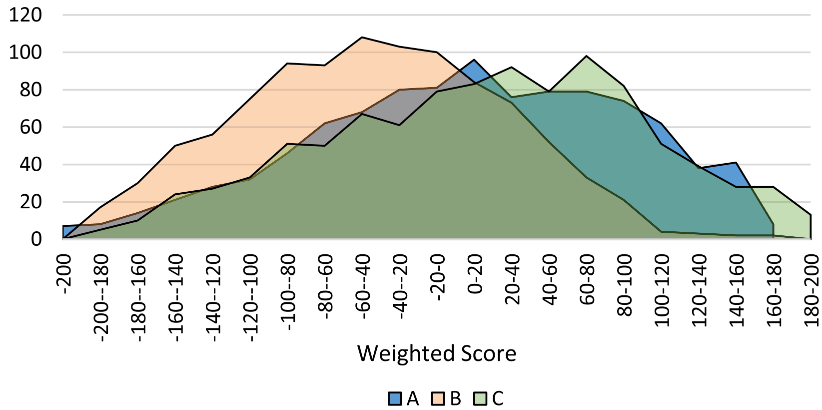

Fig. 4. Distribution of overall scores weighted scores of alternative A, B and C, based on 1000 Monte Carlo runs.

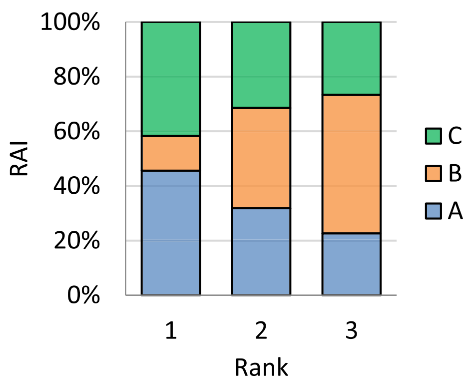

Fig. 5. Rank acceptability indices for alternatives A, B and C, based on 1000 Monte Carlo runs. The x-axis shows the rank position, and the y-axis, the rank acceptability index for each alternative (A, B and C) for every rank.

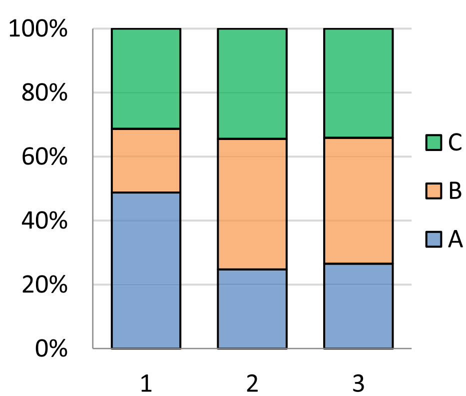

Fig. 6. Rank acceptability indices for alternatives A, B and C with doubled standard deviations in environmental indicator X, based on 1000 Monte Carlo runs.

Outranking brings benefits to the analysis in the sense that it responds to changes in mutual differences (and uncertainty), limits compensation and is compatible with importance weights. However, the fact that the assessment is context-*dependent*, makes the analysis vulnerable to instances of rank reversal when the set of alternatives changes. Rank reversal is an issue of much debate in the literature, between the proponents of descriptive versus normative approaches (Norris, 2001; Prado et al., 2012; Vargas, 1994). There is no verdict in the matter, rather it is something to take into account when applying descriptive type methods. To avoid this issue, it is of utmost importance that the analysts does a filtering of alternatives so to exclude those that are dominated entirely across environmental indicators.

Although there are studies that apply SMAA to comparative LCA studies, these represent a minority of cases. Most studies in the LCA field when aggregating apply a weighted sum (Eq. (1)). Future work in this area involves making these algorithms more accessible and transparent to broad practitioners in the field, so that aggregation approaches can move beyond a weighted sum. Practitioners in the field may not necessarily engage in active decision analysis method development, but still need to be aware of the methodological implications of aggregation and weighting. Part of the dissemination and accessibility efforts, also involves improving the visualization of results so that information is easier to understand without losing value. Moving from discrete to stochastic assessment challenges current state of practice, communication, methods, expertise and tools. Parallel to the dissemination and socialization efforts, it is important to continue to refine the methodology. Aspects to take a further look into include definition of preference thresholds given distinct probability distributions, impact category selection, addition of qualitative criteria for which there may not be impact categories yet (such as microplastics), and support to identify areas for improvement in comparative assessments (Ravikumar et al., 2018). In essence, as LCA type studies become more important in policy and industry decisions, our ability to interpret and communicate results needs to move beyond fragmented information so that these environmental studies can provide actionable insight.

Given the current tendency to consider a broader range of environmental concerns and even expand decision support to social and economic aspects (Cucurachi and Suh, 2015; Zamagni, 2012), it is important to move beyond fragmented or single issue results. Current environmental assessments such as in LCA typically deal with a dozen of environmental indicators and yet those of climate change dominate discussions. In part this is because of the relevance of climate change policy, but also because our interpretation limitations with regards to complex decision making. Aggregation can provide additional insight by including the preferences of decision makers directly into the results. In doing so, it is important that analysts are aware of the methodological implications of an aggregation method.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.envsoft.2018.08.021.

References

Aboussouan, L., Saft, R.J., Schonnenbeck, M., Hauschild, M.Z., Delbeke, K., Struijs, J., Russell, a, de Haes, H.U., Atherton, J., van Tilborg, W., Karman, C., Korenromp, R., Sap, G., Baukloh, a, Dubreuil, a, Adams, W., Heijungs, R., Jolliet, O., de Koning, a, Chapman, P., Ligthart, T., van de Meent, D., Kuyper, J., van der Loos, R., Eikelboom,

R., Verdonck, F., 2004. Declaration of Apeldoorn on LCIA of non-ferro metals. Results of a workshop by a group of LCA specialists, held in Apeldoorn, NL. SETAC Globe 5, 46–47.
Arvesen, A., Luderer, G., Pehl, M., Bodirsky, B.L., Hertwich, E.G., 2018. Deriving life cycle assessment coefficients for application in integrated assessment modelling. Environ. Model. Softw. 99, 111–125. https://doi.org/10.1016/j.envsoft.2017.09.010.
Basson, L., Petrie, J.G., 2007. An integrated approach for the consideration of uncertainty in decision making supported by Life Cycle Assessment. Environ. Model. Softw. 22, 167–176. https://doi.org/10.1016/j.envsoft.2005.07.026.
Behzadian, M., Kazemzadeh, R.B., Albadvi, a., Aghdasi, M., 2010. PROMETHEE: a comprehensive literature review on methodologies and applications. Eur. J. Oper. Res. 200, 198–215. https://doi.org/10.1016/j.ejor.2009.01.021.
Benoit, V., Rousseaux, P., 2003. Aid for aggregating the impacts in Life Cycle assessment. Int. J. Life Cycle Assess. 8, 74–82. https://doi.org/10.1007/BF02978430.
Brüggemann, R., Patil, G.P., 2011. Ranking and Prioritization for Multi-indicator Systems. Springer Verlag.
Brüggemann, R., Voigt, K., Restrepo, G., Simon, U., 2008. The concept of stability fields and hot spots in ranking of environmental chemicals. Environ. Model. Softw. 23, 1000–1012. https://doi.org/10.1016/j.envsoft.2007.11.001.
Canis, L., Linkov, I., Seager, T.P., 2010. Application of stochastic multiattribute analysis to assessment of single walled carbon nanotube synthesis processes. Environ. Sci. Technol. 44, 8704–8711. https://doi.org/10.1021/es102117k.
Castellani, V., Sala, S., Benini, L., 2016. Hotspots analysis and critical interpretation of food life cycle assessment studies for selecting eco-innovation options and for policy support. J. Clean. Prod. 1–13. https://doi.org/10.1016/j.jclepro.2016.05.078.
Cellura, M., Di Gangi, A., Longo, S., Orioli, A., 2013. An Italian input–output model for the assessment of energy and environmental benefits arising from retrofit actions of buildings. Energy Build. 62, 97–106. https://doi.org/10.1016/j.enbuild.2013.02.056.
Cinelli, M., Coles, S.R., Kirwan, K., 2014. Analysis of the potentials of multi criteria decision analysis methods to conduct sustainability assessment. Ecol. Indic. 46, 138–148. https://doi.org/10.1016/j.ecolind.2014.06.011.
Cucurachi, S., Seager, T.P., Prado, V., 2017. Normalization in comparative life cycle assessment to support environmental decision making. J. Ind. Ecol. 21, 242–243. https://doi.org/10.1111/jiec.12549.
Cucurachi, S., Suh, S., 2015. A moonshot for sustainability assessment. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.5b02960.
Domingues, A.R., Marques, P., Garcia, R., Freire, F., Dias, L.C., 2015. Applying multi-criteria decision analysis to the life-cycle assessment of vehicles. J. Clean. Prod. 107, 749–759. https://doi.org/10.1016/j.jclepro.2015.05.086.
Figueira, J.R., Mousseau, V., Roy, B., 2016. ELECTRE methods. In: Greco, S., Ehrgott, M., Figueira, J.R. (Eds.), Multiple Criteria Decision Analysis. State of the Art Surveys. Springer, New York, pp. 155–185. https://doi.org/10.1007/978-1-4939-3094-4_5.
Finnveden, G., Hauschild, M.Z., Ekvall, T., Guinée, J., Heijungs, R., Hellweg, S., Koehler, A., Pennington, D., Suh, S., 2009. Recent developments in life cycle assessment. J. Environ. Manage. 91, 1–21. https://doi.org/10.1016/j.jenvman.2009.06.018.
Fischer, G., 1995. Range sensitivity of attribute weights in multiattribute value models. Organ. Behav. Hum. Decis. Process. 62, 252–266.
Gagnon, L., Bélanger, C., Uchiyama, Y., 2002. Life-cycle assessment of electricity generation options: the status of research in year 2001. Energy Pol. 30, 1267–1278. https://doi.org/10.1016/S0301-4215(02)00088-5.
Groen, E.A., Heijungs, R., Bokkers, E.A.M., de Boer, I.J.M., 2014. Methods for uncertainty propagation in life cycle assessment. Environ. Model. Softw. 62, 316–325. https://doi.org/10.1016/j.envsoft.2014.10.006.
Hammond, J., Keeney, R.L., Raiffa, H., 1999. Smart Choices. Harvard Business School Press, Boston, MA.
Heijungs, R., Guinée, J., Kleijn, R., Rovers, V., 2007. LCA methodology bias in Normalization : causes, consequences, detection and remedies *. Int. J. Life Cycle Assess. 12, 211–216.
Heijungs, R., Henriksson, P., Guinée, J., 2016. Measures of difference and significance in the era of computer simulations, meta-analysis, and big data. Entropy 18, 361. https://doi.org/10.3390/e18100361.
Hellweg, S., Mila i Canals, L., 2014. Emerging approaches, challenges and opportunities in life cycle assessment. Science (80-.) 344, 1109–1113. https://doi.org/10.1126/science.1248361.
Hertwich, E.G., Hammitt, J.K., 2001. A decision-analytic framework for impact assessment part I: LCA and decision analysis. Int. J. Life Cycle Assess. 6, 5–12. https://doi.org/10.1007/bf02977588.
Hertwich, E.G., Hammitt, J.K., Pease, W.S., 2000. A theoretical foundation for life-cycle assessment. J. Ind. Ecol. 4, 13–28. https://doi.org/10.1162/108819800569267.
Janssen, R., 1992. Multiobjective Decision Support for Environmental Management. Springer, Netherlands.
Keeney, R.L., 2002. Common mistakes in making value trade-offs. Oper. Res. 50, 935–945. https://doi.org/10.1287/opre.50.6.935.357.
Kiker, G.A., Bridges, T.S., Varghese, A., Seager, T.P., Linkov, I., 2005. Application of multicriteria decision analysis in environmental decision making. Integr. Environ. Assess. Manag. 1, 95. https://doi.org/10.1897/IEAM_2004a-015.1.
Kim, J., Yang, Y., Bae, J., Suh, S., 2013. The importance of normalization references in interpreting life cycle assessment results. J. Ind. Ecol. 17, 385–395. https://doi.org/10.1111/j.1530-9290.2012.00535.x.
Lahdelma, R., Salminen, P., 2001. SMAA-2 : stochastic multicriteria acceptability analysis for group decision making. Oper. Res. 49, 444–454. https://doi.org/10.2307/3088639.
Laurin, L., Amor, B., Bachmann, T.M., Bare, J., Koffler, C., Genest, S., Preiss, P., Pierce, J., Satterfield, B., Vigon, B., 2016. Life cycle assessment capacity roadmap (section 1): decision-making support using LCA. Int. J. Life Cycle Assess. 21, 443–447. https://doi.org/10.1007/s11367-016-1031-y.
Lautier, A., Rosenbaum, R.K., Margni, M., Bare, J., Roy, P.-O., Deschênes, L., 2010. Development of normalization factors for Canada and the United States and comparison with European factors. Sci. Total Environ. 409, 33–42. https://doi.org/10.1016/j.scitotenv.2010.09.016.
Matarazzo, A., Clasadonte, M.T., Ingrao, C., Zerbo, A., 2013. Criteria interaction modelling in the framework of Lca analysis. Int. J. Eng. Res. Appl. 3, 523–530.
Mendoza Beltran, A., Prado, V., Font Vivanco, D., Henriksson, P.J.G., Guinée, J.B., Heijungs, R., 2018. Quantified uncertainties in comparative life cycle assessment: what can Be concluded? Environ. Sci. Technol. https://doi.org/10.1021/acs.est.7b06365. acs.est.7b06365.
Munda, G., 2008. Social Multi-criteria Evaluation for a Sustainable Economy. Springer Berlin Heidelberg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-73703-2.
Munda, G., Nardo, M., 2009. Noncompensatory/nonlinear composite indicators for ranking countries: a defensible setting. Appl. Econ. 14, 1513–1523. https://doi.org/10.1080/00036840601019364.
Norris, G. a., 2001. The requirement for congruence in normalization. Int. J. Life Cycle Assess. 6, 85–88. https://doi.org/10.1007/BF02977843.
Pauliuk, S., Arvesen, A., Stadler, K., Hertwich, E.G., 2017. Industrial ecology in integrated assessment models. Nat. Clim. Chang. 7, 13–20. https://doi.org/10.1038/nclimate3148.
Pizzol, M., Christensen, P., Schmidt, J., Thomsen, M., 2011. Impacts of "metals" on human health: a comparison between nine different methodologies for Life Cycle Impact Assessment (LCIA). J. Clean. Prod. 19, 646–656. https://doi.org/10.1016/j.jclepro.2010.05.007.
Polatidis, H., Haralambopoulos, D.A., Vreeker, R., 2006. Selecting an appropriate multi-criteria decision analysis technique for renewable energy planning. Energy Sources, Part B 1, 181–193. https://doi.org/10.1080/009083190881607.
Pollesch, N., Dale, V.H., 2015. Applications of aggregation theory to sustainability assessment. Ecol. Econ. 114, 117–127. https://doi.org/10.1016/j.ecolecon.2015.03.011.
Pollesch, N.L., Dale, V.H., 2016. Normalization in sustainability assessment: methods and implications. Ecol. Econ. 130, 195–208. https://doi.org/10.1016/j.ecolecon.2016.06.018.
Prado-Lopez, V., Seager, T.P., Chester, M., Laurin, L., Bernardo, M., Tylock, S., 2014. Stochastic multi-attribute analysis (SMAA) as an interpretation method for comparative life-cycle assessment (LCA). Int. J. Life Cycle Assess. 19, 405–416. https://doi.org/10.1007/s11367-013-0641-x.
Prado-Lopez, V., Wender, B. a, Seager, T.P., Laurin, L., Chester, M., 2015. Tradeoff evaluation improves a photovoltaic case study. J. Ind. Ecol. 0, 1–9. https://doi.org/10.1111/jiec.12292.
Prado, V., Rogers, K., Seager, T.P., 2012. Integration of MCDA tools in valuation of comparative life cycle assessment. In: Curran, M.A. (Ed.), Life Cycle Assessment Handbook: a Guide for Environmentally Sustainable Products. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 413–431. https://doi.org/10.1002/9781118528372.
Prado, V., Wender, B.A., Seager, T.P., 2017. Interpretation of comparative LCAs: external normalization and a method of mutual differences. Int. J. Life Cycle Assess. 22, 2018–2029. https://doi.org/10.1007/s11367-017-1281-3.
Rajagopalan, N., Venditti, R., Kelley, S., Daystar, J., 2016. Multi-attribute uncertainty analysis of the life cycle of lignocellulosic feedstock for biofuel production. Biofuels, Bioprod. Biorefining. https://doi.org/10.1002/bbb.1737.
Ravikumar, D., Seager, T.P., Cucurachi, S., Prado, V., Mutel, C.L., 2018. Novel method of sensitivity analysis improves the prioritization of research in anticipatory life cycle assessment of emerging technologies. Environ. Sci. Technol. 52, 6534–6543. https://doi.org/10.1021/acs.est.7b04517.
Reichert, P., Borsuk, M., Hostmann, M., Schweizer, S., Spörri, C., Tockner, K., Truffer, B., 2007. Concepts of decision support for river rehabilitation. Environ. Model. Softw. 22, 188–201. https://doi.org/10.1016/j.envsoft.2005.07.017.
Riabacke, M., Danielson, M., Ekenberg, L., 2012. State-of-the-art prescriptive criteria weight elicitation. Adv. Decis. Sci. 2012. https://doi.org/10.1155/2012/276584.
Rogers, K., 2008. Environmental Decision-making Using Life Cycle Impact Assessment and Stochastic Muti-attribute Decision Analysis: a Case Study on Alternative Transportation Fuels. Purdue University.
Rogers, K., Seager, T.P., 2009. Environmental decision-making using life cycle impact assessment and stochastic multiattribute decision analysis: a case study on alternative transportation fuels. Environ. Sci. Technol. 43, 1718–1723. https://doi.org/10.1021/es801123h.
Rogers, M., Bruen, M., 1998. Choosing realistic values of indifference, preference and veto thresholds for use with environmental criteria within ELECTRE. Eur. J. Oper. Res. 107, 542–551. https://doi.org/10.1016/S0377-2217(97)00175-6.
Rowley, H.V., Peters, G.M., Lundie, S., Moore, S.J., 2012. Aggregating sustainability indicators: beyond the weighted sum. J. Environ. Manage. 111, 24–33. https://doi.org/10.1016/j.jenvman.2012.05.004.
Roy, B., 1985. Méthodologie multicritère d'aide à la décision. Economica, Paris.
Stewart, T.J., 2008. Robustness analysis and MCDA. Eur. Work. Gr. Mult. Criteria Decis. Aiding 3.
Tan, R.R., 2005. Rule-based life cycle impact assessment using modified rough set induction methodology. Environ. Model. Softw. 20, 509–513. https://doi.org/10.1016/j.envsoft.2004.08.005.
Tervonen, T., Lahdelma, R., 2007. Implementing stochastic multicriteria acceptability analysis. Eur. J. Oper. Res. 178, 500–513. https://doi.org/10.1016/j.ejor.2005.12.037.
Tsoukiàs, A., 2008. From decision theory to decision aiding methodology. Eur. J. Oper. Res. 187, 138–161. https://doi.org/10.1016/j.ejor.2007.02.039.
Tylock, S.M., Seager, T.P., Snell, J., Bennett, E.R., Sweet, D., 2012. Energy management under policy and technology uncertainty. Energy Pol. 47, 156–163. https://doi.org/

10.1016/j.enpol.2012.04.040.
Vargas, L.G., 1994. Theory and Methodology Reply to Schenkerman â€™ s Avoiding Rank Reversal in AHP Decision Support Models 74. pp. 420–425.
Wender, B., Foley, R.W., Prado-lopez, V., Ravikumar, D., Eisenberg, D.A., Hottle, T.A., Sadowski, J., Flanagan, W.P., Fisher, A., Laurin, L., Bates, M.E., Linkov, I., Seager, T.P., Fraser, M.P., Guston, D.H., 2014. Illustrating anticipatory life cycle assessment for emerging photovoltaic technologies. Environ. Sci. Technol. 2014, 10531–10538.
White, P., Carty, M., 2010. Reducing bias through process inventory dataset normalization. Int. J. Life Cycle Assess. 15, 994–1013. https://doi.org/10.1007/s11367-010-0215-0.
Zamagni, A., 2012. Life cycle sustainability assessment. Int. J. Life Cycle Assess. 17, 373–376. https://doi.org/10.1007/s11367-012-0389-8.
Zhu, F., Zhong, P. an, Sun, Y., 2018. Multi-criteria group decision making under uncertainty: application in reservoir flood control operation. Environ. Model. Softw. 100, 236–251. https://doi.org/10.1016/j.envsoft.2017.11.032.